

Nonnormality of Stoneham constants

David H. Bailey* Jonathan M. Borwein†

January 14, 2012

Abstract

Previous studies have established that Stoneham's constant $\alpha_{2,3} = \sum_{n \geq 1} 1/(3^n 2^{3^n})$ is 2-normal, or, in other words, every m -long string of binary digits appears in the binary expansion of $\alpha_{2,3}$ with precisely the expected limiting frequency $1/2^m$. A more recent finding is that this constant is provably *not* 6-normal. In this note we address the more general class of Stoneham constants $\alpha_{b,c} = \sum_{n \geq 1} 1/(c^n b^{c^n})$, for coprime integers $b \geq 2$ and $c \geq 2$. It has been proven that $\alpha_{b,c}$ is b -normal, but *not* bc -normal. Here we extend this finding by showing that $\alpha_{b,c}$ is *not* B -normal, where $B = b^p c^q r$, for integers b and c as above, $p, q, r \geq 1$, neither b nor c divide r , and the condition $D = c^{q/p} r^{1/p} / b^{c-1} < 1$ is satisfied. It is not known whether or not this is a complete catalog of bases to which $\alpha_{b,c}$ is nonnormal. We also show that the sum of two B -nonnormal Stoneham constants as defined above is B -nonnormal.

1 Introduction

The question of whether (and why) the digits of well-known constants of mathematics are statistically random in some sense has fascinated mathematicians from the dawn of history. Indeed, one prime motivation in computing and analyzing digits of π is

*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, DHBailey@lbl.gov. Supported in part by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC02-05CH11231.

†Laureate Professor and Director Centre for Computer Assisted Research Mathematics and its Applications (CARMA), University of Newcastle, Callaghan, NSW 2308, Australia. Distinguished Professor, King Abdulaziz University, Jeddah 80200, Saudi Arabia. Email: jonathan.borwein@newcastle.edu.au.

Table 1: Digit counts in the first trillion hexadecimal (base-16) digits of π . Note that deviations from the average value 62,500,000,000 occur only after the first six digits, as expected from the central limit theorem.

Hex Digit	Occurrences
0	62499881108
1	62500212206
2	62499924780
3	62500188844
4	62499807368
5	62500007205
6	62499925426
7	62499878794
8	62500216752
9	62500120671
A	62500266095
B	62499955595
C	62500188610
D	62499613666
E	62499875079
F	62499937801
Total	1000000000000

to explore the age-old question of whether and why these digits appear “random.” The first computation on ENIAC in 1949 of π to 2037 decimal places was proposed by John von Neumann so as to shed some light on the distribution of π (and of e) [8, pg. 277–281].

Since then, numerous computer-based statistical checks of the digits of π , for instance, so far have failed to disclose any deviation from reasonable statistical norms. See, for instance, Table 1, which presents the counts of individual hexadecimal digits among the first trillion hex digits, as obtained by Yasumasa Kanada. By contrast, the early computations did reveal provable abnormalities in the behavior of e [11, §11.2]. Figures 1 and 2 compare a random walk based on the base-4 digits of a pseudorandomly generated real, compared with with a walk based on the binary digits of π . In each case the “walk” is defined as follows: move up or down one unit, depending on whether the binary digit in the given odd-numbered position is a zero or a one, then move left or right depending on whether the next binary digit in an even-numbered position is a zero or a one. For additional details and analysis, the

reader should consult [2].

In the following, we say a real constant α is b -normal if, given the positive integer $b \geq 2$, every m -long string of base- b digits appears in the base- b expansion of α with precisely the expected limiting frequency $1/b^m$. It is a well-established albeit counter-intuitive fact that given an integer $b \geq 2$, almost all real numbers, in the measure theory sense, are b -normal. What's more, almost all real numbers are b -normal simultaneously for all positive integer bases (a property known as “absolutely normal”).

The present perplexing situation is that whereas almost all real numbers are normal for all integer bases $b \geq 2$, it has been frustratingly difficult to exhibit explicit examples, even of reals that are normal just to a single given base b . The first constant to be proven 10-normal is the Champernowne number, namely the constant $0.12345678910111213141516\dots$, produced by concatenating the decimal representation of all positive integers in order. Some additional results of this sort were established in the 1940s by Copeland and Erdős [14].

The situation with regards to other, more “natural” constants of mathematics remains singularly grim. Normality proofs are not available for any well-known constant such as $\pi, e, \log 2, \sqrt{2}$. We do not even know, say, that a 1 appears $1/2$ of the time, in the limit, in the binary expansion of $\sqrt{2}$ (although it certainly appears to, from extensive empirical analysis). For that matter, it is widely believed that *every* irrational algebraic number (i.e., every irrational root of an algebraic polynomial with integer coefficients) is b -normal to all positive integer bases, but there is no proof.

Recently the present authors, together with Richard Crandall and Carl Pomerance, proved the following: If a real y has algebraic degree $D > 1$, then the number $\#(|y|, N)$ of 1-bits in the binary expansion of $|y|$ through bit position N satisfies

$$\#(|y|, N) > CN^{1/D} \tag{1}$$

for a positive number C (depending on y) and all sufficiently large N [3]. For example, there must be at least \sqrt{N} 1-bits in the first N bits in the binary expansion of $\sqrt{2}$, in the limit. A related and more refined result has been obtained by Hajime Kaneko of Kyoto University in Japan. He obtained the bound in $C(\log N)^{3/2}/[(\log(6D))^{1/2}(\log \log N)^{1/2}]$ and extended his results to a very general class of bases and algebraic irrationals [16]. However, each of these results falls far short of establishing b -normality for any irrational algebraic in any base b , even in the single-digit sense.

It is known that whenever α is b -normal, then so is $r\alpha$ and $r + \alpha$ for any nonzero positive rational r [9, pg. 165–166]. It is also easy to see that if there is a positive integer n such that integers $a \geq 2$ and $b \geq 2$ satisfy $a = b^n$, then any real constant

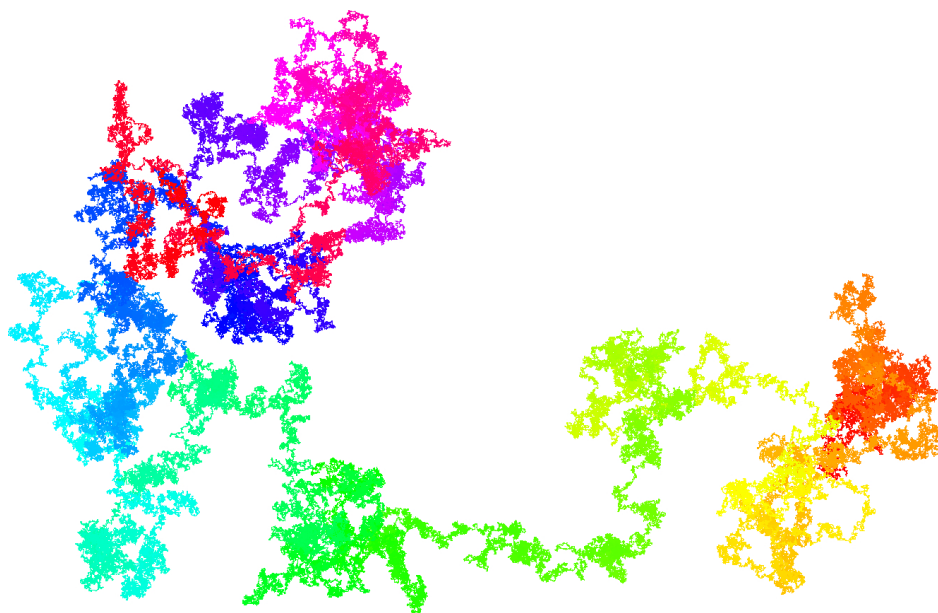


Figure 1: A uniform pseudorandom walk.

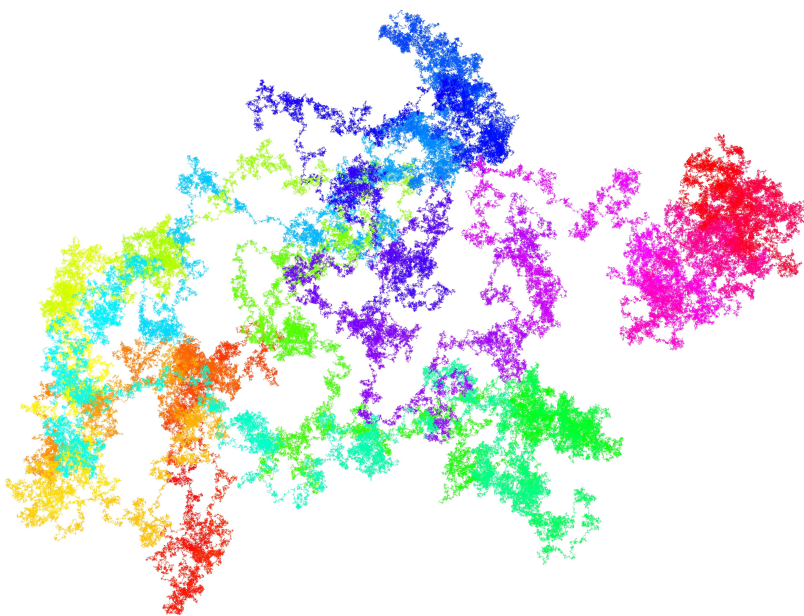


Figure 2: A random walk on the first two billion bits of π (normal?).

that is a -normal is also b -normal. Recently Hertling proved an interesting converse: If there is no such n , then there are an uncountable number of counterexamples, namely constants that are a -normal but not b normal [15].

Moving in the other direction, Greg Martin has succeeded in constructing an absolutely nonnormal number, namely one which fails to be b -normal for any integer base $b \geq 2$ [18]. We shall discuss b -nonnormal constants to a fixed base b below.

Numerous other known results relating to normality are summarized by Khoshnevisan [17] and also in [2].

2 A recent normality result

In 2002, one of the present authors (Bailey) and Richard Crandall showed that given a real number r in $[0, 1)$, with r_k denoting the k -th binary digit of r , the real number

$$\alpha_{2,3}(r) : = \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k + r_k}} \quad (2)$$

is 2-normal. It can be seen that if $r \neq s$, then $\alpha_{2,3}(r) \neq \alpha_{2,3}(s)$, so that these constants are all distinct. Since r can range over the unit interval, this class of constants is uncountable. So, for example, the constant $\alpha_{2,3} = \alpha_{2,3}(0) = \sum_{k \geq 1} 1/(3^k 2^{3^k}) = 0.0418836808315030 \dots$ is provably 2-normal (this special case was proven by Stoneham in 1973 [19]). A similar result applies if 2 and 3 in formula (2) are replaced by any pair of coprime integers (b, c) with $b \geq 2$ and $c \geq 2$ [4].

More recently, Bailey and Michal Misieurwicz were able to establish 2-normality of $\alpha_{2,3}$ by a simpler argument, by utilizing a “hot spot” lemma proven using ergodic theory methods [5]. In [1], this proof was extended to the more general case $\alpha_{b,c}$, although the result itself was established in the 2002 Bailey-Crandall paper. We reprise this proof below, preceded by a statement of the “hot spot lemma” from [5].

Let $A(\alpha, y, n, m)$ denote the count of occurrences where the m -long binary string y is found to start at position p in the base- b expansion of α , where $1 \leq p \leq n$.

Lemma 1 (“Hot Spot” Lemma): *If x is not b -normal, then there is some $y \in [0, 1)$ with the property*

$$\liminf_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{b^m A(x, y, n, m)}{n} = \infty. \quad (3)$$

Conversely, if for all $y \in [0, 1)$,

$$\liminf_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{b^m A(x, y, n, m)}{n} < \infty, \quad (4)$$

then x is b -normal.

Note that Lemma 1 implies that if a real constant α is not b -normal, then there must exist some interval $[r_1, s_1)$ with the property that successive shifts of the base- b expansion of α visit $[r_1, s_1)$ ten times more frequently, in the limit, relative to its length $s_1 - r_1$; there must be another interval $[r_2, s_2)$ that is visited 100 times more often relative to its length; there must be a third interval $[r_3, s_3)$ that is visited 1,000 times more often relative to its length; etc. Furthermore, there exists at least one real number y (a “hot spot”) such that sufficiently small neighborhoods of y are visited too often by an arbitrarily large factor, relative to the lengths of these neighborhoods. On the other hand, if it can be established that no subinterval of the unit interval is visited 1,000 times (for instance) more often in the limit relative to its length, then this suffices to prove that the constant in question is b -normal (and thus that each subinterval is visited with precisely the correct frequency, in the limit, relative to the size of the subinterval).

Theorem 1 *For every coprime pair of integers (b, c) with $b \geq 2$ and $c \geq 2$, the constant $\alpha_{b,c} = \sum_{m \geq 1} 1/(c^m b^{c^m})$ is b -normal.*

Proof: We can write the the fraction immediately following position n in the base- b expansion of $\alpha_{b,c}$ as:

$$b^n \alpha_{b,c} \bmod 1 = \left(\sum_{m=1}^{\infty} \frac{b^{n-c^m} \bmod c^m}{c^m} \right) \bmod 1 \quad (5)$$

$$= \left(\sum_{m=1}^{\lfloor \log_c n \rfloor} \frac{b^{n-c^m} \bmod c^m}{c^m} \right) \bmod 1 + \sum_{m=\lfloor \log_c n \rfloor + 1}^{\infty} \frac{b^{n-c^m}}{c^m}. \quad (6)$$

Note that the first expression can be generated by means of the recursion $z_0 = 0$ and, for $n \geq 1$, $z_n = (bz_{n-1} + r_n) \bmod 1$, where $r_n = 1/n$ if $n = c^k$ for some integer k , and zero otherwise. For example, consider the case $b = 3$ and $c = 4$. The first few

members of the z sequence are given as follows:

$$\begin{aligned}
& 0, 0, 0, \\
& \frac{1}{4}, \frac{3}{4}, \quad (\text{repeated 6 times}) \\
& \frac{5}{16}, \frac{15}{16}, \frac{13}{16}, \frac{7}{16}, \quad (\text{repeated 12 times}), \\
& \frac{21}{64}, \frac{63}{64}, \frac{61}{64}, \frac{55}{64}, \frac{37}{64}, \frac{47}{64}, \frac{13}{64}, \frac{39}{64}, \frac{53}{64}, \frac{31}{64}, \frac{29}{64}, \frac{23}{64}, \frac{5}{64}, \frac{15}{64}, \frac{45}{64}, \frac{7}{64}, \\
& (\text{repeated 12 times}), \text{ etc.}
\end{aligned} \tag{7}$$

Note here that the fraction $1/2$ is omitted in the first set, the fractions $1/8, 3/8, 5/8, 7/8$ are omitted in the second set, and the fractions with 32 in the denominators are omitted in the third set. Nonetheless, this critical property holds, both in this particular case and in general, so long as $b \geq 2$ and $c \geq 2$ are coprime [4]: if $n < c^{p+1}$ then z_n is a multiple of $1/c^p$, and furthermore the set $(z_k, 1 \leq k \leq n)$ contains at most t repetitions of any particular value, where the integer t depends only on (b, c) . For the case $(2, 3)$, the repetition factor $t = 3$. For the case $(3, 4)$, $t = 12$.

These fractions (z_k) constitute an accurate set of approximations to the sequence $b^n \alpha_{b,c} \bmod 1$ of shifted fractions of $\alpha_{b,c}$, since (z_k) generates the first term of (6). In fact, by examining (6) it can be readily seen that for all (b, c) as above and all $n \geq c$,

$$|b^n \alpha_{b,c} \bmod 1 - z_n| < \frac{1}{9n} \tag{8}$$

(and in most cases is much smaller than this).

To establish that $\alpha_{b,c}$ is b -normal via Lemma 1, we seek an upper bound for $b^m A(\alpha_{b,c}, y, n, m)/n$, good for all $y \in [0, 1)$ and all $m \geq 1$. A binary sequence y out to some length m , translated to a subset of the real unit interval, can be written as $[r, s)$, where $r = 0.y_1 y_2 y_3 \dots y_m$, and s is the next largest base- b fraction of length m , so that $s - r = b^{-m}$. Observe that the count $A(\alpha_{b,c}, y, n, m)$ is equal to the number of those j between 0 and $n - 1$ for which $b^j \alpha_{b,c} \bmod 1 \in [r, s)$. Also observe, in view of (8), that if $b^j \alpha \bmod 1 \in [r, s)$, then $z_j \in [r - 1/(9j), s + 1/(9j))$.

Let n be any integer greater than b^{2m} , and let c^p denote the largest power of c less than or equal to n , so that $c^p \leq n < c^{p+1}$. Now note that for $j \geq b^m$, we have $[r - 1/(9j), s + 1/(9j)) \subset [r - b^{-m-1}, s + b^{-m-1})$. Since the length of this latter interval is no greater than $2b^{-m}$, the number of multiples of $1/c^p$ that it contains cannot exceed $\lfloor 2c^p b^{-m} \rfloor + 1$. Thus there can be at most t times this many j 's less

than n for which $z_j \in [r - b^{-m-1}, s + b^{-m-1}]$. Therefore we can write

$$\begin{aligned} \frac{b^m A(\alpha_{b,c}, y, n, m)}{n} &= \frac{b^m \#_{0 \leq j < n} (b^j \alpha_{b,c} \bmod 1 \in [r, s])}{n} \\ &\leq \frac{b^m [b^m + \#_{b^m \leq j < n} (z_j \in [r - b^{-m-1}, s + b^{-m-1}])]}{n} \\ &\leq \frac{b^m [b^m + t(2c^p b^{-m} + 1)]}{n} < 2t + 2, \end{aligned} \quad (9)$$

where t is the fixed repetition factor for (b, c) , mentioned above. For a fixed pair of integers (b, c) , we have shown that for all $y \in [0, 1]$ and all $m > 0$,

$$\limsup_{n \rightarrow \infty} \frac{b^m A(\alpha_{b,c}, y, n, m)}{n} \leq 2t + 2, \quad (10)$$

so by Lemma 1, $\alpha_{b,c}$ is b -normal. **QED**

This result was first proven by Bailey and Crandall in 2002 [4]. The proof above, which utilizes the hot spot lemma, appeared in [1].

3 A general nonnormality result

By Theorem 1, the Stoneham constant $\alpha_{2,3} = \sum_{k \geq 0} 1/(3^k 2^{3^k})$ is 2-normal. Almost as interesting is the fact that $\alpha_{2,3}$ is *not* 6-normal, and, more generally, $\alpha_{b,c}$ (where b and c are restricted as above) is *not* bc -normal. This was first demonstrated in [1]. We will sketch the proof here for the case $\alpha_{2,3}$. The complete proof of an even more general result is presented in the Appendix.

First note that the digits immediately following position n in the base-6 expansion of $\alpha_{2,3}$ can be obtained by computing $6^n \alpha_{2,3} \bmod 1$, which can be written as

$$6^n \alpha_{2,3} \bmod 1 = \left(\sum_{m=1}^{\lfloor \log_3 n \rfloor} 3^{n-m} 2^{n-3^m} \right) \bmod 1 + \sum_{m=\lfloor \log_3 n \rfloor + 1}^{\infty} 3^{n-m} 2^{n-3^m}. \quad (11)$$

Note that the first portion of this expression is *zero*, since all terms of the summation are integers. That leaves the second expression.

Consider the case when $n = 3^m$, where $m \geq 1$ is an integer, and examine just the first term of the second summation. We see that this expression is

$$3^{3^m - (m+1)} 2^{3^m - 3^{m+1}} = 3^{3^m - m - 1} 2^{-2 \cdot 3^m} = (3/4)^{3^m} / 3^{m+1}. \quad (12)$$

m	3^m	Z_m
1	3	1
2	9	3
3	27	6
4	81	16
5	243	42
6	729	121
7	2187	356
8	6561	1058
9	19683	3166
10	59049	9487

Table 3: Counts Z_m of consecutive zeroes immediately following position 3^m in the base-6 expansion of $\alpha_{2,3}$.

It is worth pointing out here that in the parlance of Lemma 1, zero is a “hot spot” for the base-6 expansion of $\alpha_{2,3}$. This is because all sufficiently small neighborhoods of zero are visited too often, by an arbitrarily large factor, in a subsequence of the shifted fractions of its base-6 expansion.

Here is a generalization of the result for general Stoneham constants $\alpha_{b,c}$:

Theorem 2 *Given coprime integers $b \geq 2$ and $c \geq 2$, and integers $p, q, r \geq 1$, with neither b nor c dividing r , let $B = b^p c^q r$. Assume that the condition $D = c^{q/p} r^{1/p} / b^{c-1} < 1$ is satisfied. Then the constant $\alpha_{b,c} = \sum_{k \geq 0} 1/(c^k b^{c^k})$ is B -nonnormal.*

Proof: See the Appendix for the proof of this new result.

QED

The following less general result than Theorem 2 first appeared in [1]:

Corollary 1 *Given coprime integers $b \geq 2$ and $c \geq 2$, $\alpha_{b,c}$ is bc -nonnormal.*

Proof: This is a special case of Theorem 2 where $p = q = r = 1$. It follows by checking the condition (see the hypothesis of Theorem 2) that $D = c/b^{c-1} < 1$, or, equivalently, that $\log c < (c-1) \log b$. This condition can be verified as follows. First assume that $b \geq 2$ and $c \geq 3$. In this case, the function $f(c) = \log c - (c-1) \log 2 < 0$, so that $\log c < (c-1) \log 2 \leq (c-1) \log b$. Similarly, when $b \geq 3$ and $c \geq 2$, the function $g(c) = \log c - (c-1) \log 3 < 0$, so that $\log c < (c-1) \log 3 \leq (c-1) \log b$. The remaining case $b = 2$ and $c = 2$ is not allowable, since b and c must be coprime.

Thus the key condition $c/b^{c-1} < 1$ in the hypothesis of Theorem 2 is satisfied by all allowable pairs (b, c) . Hence $\alpha_{b,c}$ is not bc -normal. **QED**

Example 1 (Normality and nonnormality in various bases) According to Theorem 1, the constant $\alpha_{2,3}$ is normal base 2, and thus is also normal in base 4, 8, 16, 32, \dots (i.e., all powers of two). According to Theorem 2, $\alpha_{2,3}$ is nonnormal base 6, 12, 24, 36, 48, 60, 72, 96, 120, 144, 168, 192, 216, 240, \dots . This list can be obtained by checking the condition $3^{q/p} r^{1/p} < 4$ for various candidate bases $B = 2^p 3^q r$, where $p, q, r \geq 1$. Note that while all integers in this list are divisible by 6, not all multiples of 6 are in the list.

However, there are many integer bases not included in either list. For example, it is not known at the present time whether or not $\alpha_{2,3}$ is 3-normal, although it appears to be. For example, statistical analysis of the first 83,736 base-3 digits of $\alpha_{2,3}$ (both single digits and 6-long strings of digits) found no deviations from reasonable statistical norms. But there is no proof of 3-normality. Similar questions remain in the more general case of $\alpha_{b,c}$, where b and c are coprime and at least two.

4 Sums of Stoneham constants

We now examine the normality or nonnormality of the sum of two Stoneham constants.

Under the hypothesis $b, c_1, c_2 \geq 2$, with (b, c_1) coprime and (b, c_2) coprime, we know from Theorem 1 that α_{b,c_1} and α_{b,c_2} are each b -normal. But it is not known at the present time whether the sum $\alpha_{b,c_1} + \alpha_{b,c_2}$ is b -normal. However, the sum of two such constants that individually are B -nonnormal, for some base B as given in the hypothesis of Theorem 2, is also B -nonnormal:

Theorem 3 *Let α_{b_1,c_1} and α_{b_2,c_2} be two Stoneham constants satisfying the conditions of Theorem 2 to be B -nonnormal: $b_1 \geq 2$ and $c_1 \geq 2$ are coprime; $B = b_1^{p_1} c_1^{q_1} r_1$ for integers $p_1, q_1, r_1 \geq 1$ with neither b_1 nor c_1 dividing r_1 ; and $D_1 = c_1^{q_1/p_1} r_1^{1/p_1} / b_1^{c_1-1} < 1$ (with similar conditions on b_2, c_2, p_2, q_2, r_2 and D_2). Then $\alpha_{b_1,c_1} + \alpha_{b_2,c_2}$ is B -nonnormal.*

Proof: See the Appendix for the proof of this new result. **QED**

Example 2 (Nonnormality of sums in various bases) Consider the Stoneham constants $\alpha_{2,3}$ and $\alpha_{2,5}$. By Theorem 1, both are 2-normal. Consider base $60 = 2^p \cdot 3^q \cdot r$, where $p = 2, q = 1$ and $r = 5$. By checking the condition $3^{1/2} \cdot 5^{1/2} < 2^2$,

we verify that $\alpha_{2,3}$ is 60-nonnormal, according to Theorem 2. In a similar way, write $60 = 2^p \cdot 5^q \cdot r$, where $p = 2, q = 1$ and $r = 3$. Then by checking the condition $5^{1/2} \cdot 3^{1/2} < 2^4$, we verify that $\alpha_{2,5}$ is also 60-nonnormal. Thus, according to Theorem 3, $\alpha_{2,3} + \alpha_{2,5}$ is 60-nonnormal.

5 Conclusion

As mentioned above, under the hypothesis that integers $b \geq 2$, $c_1 \geq 2$ and $c_2 \geq 2$ are coprime, we know from Theorem 1 that α_{b,c_1} and α_{b,c_2} are each b -normal, but it is not known at the present time whether the sum $\alpha_{b,c_1} + \alpha_{b,c_2}$ is b -normal (although from substantial empirical analysis of specific cases, this appears to be true). Such a result, if it could be proven and extended, may yield a construction of an explicit computable constant that is absolutely normal, namely b -normal for all integer bases $b \geq 2$ simultaneously.

One example of an absolutely normal constant, as defined in the previous paragraph, is Chaitin's omega constant. Fix a prefix-free universal Turing machine U : (i.e., if instances $U(p)$ and $U(q)$ each halt, then neither p nor q is a prefix of the other.) Then Chaitin's omega is defined by

$$\Omega = \sum_{\{U(p) \text{ halts}\}} 2^{-|p|},$$

where $|p|$ is the length of the program p in bits. In 1994, Cristian Calude [12] demonstrated that omega is absolutely normal. Although a scheme is known to explicitly compute the value of an initial segment of Chaitin's constant (for a certain encoding of a Turing machine), fewer than 100 bits are known [13]. Another explicit construction has been given by Becher and Figueira [6]. However, unlike Chaitin's constant, while it is possible in principle to compute digits of the the Becher-Figueira constant, it is nigh impossible in practice. It transpires that Alan Turing visited this same issue many decades ago — as described in [7].

In any event, there is continuing interest in explicitly constructive real numbers that are both absolutely normal and which can be computed to high precision with reasonable effort.

References

- [1] David Bailey and Jonathan Borwein, “Normal numbers and pseudorandom generators,” *Proceedings of the Workshop on Computational and Analytical Mathematics in Honour of Jonathan Borwein’s 60th Birthday*, Springer, to appear, September 2011, available at <http://crd.lbl.gov/~dhbailey/dhbpapers/normal-pseudo.pdf>.
- [2] David H. Bailey, Jonathan M. Borwein, Cristian S. Calude, Michael J. Dinneen, Monica Dumitrescu and Alex Yee, “An empirical approach to the normality of pi,” manuscript, 9 Jan 2012, available at <http://crd.lbl.gov/~dhbailey/dhbpapers/normality.pdf>.
- [3] David H. Bailey, Jonathan M. Borwein, Richard E. Crandall and Carl Pomerance, “On the binary expansions of algebraic numbers,” *Journal of Number Theory Bordeaux*, vol. 16 (2004), pg. 487–518.
- [4] David H. Bailey and Richard E. Crandall, “Random generators and normal numbers,” *Experimental Mathematics*, vol. 11 (2002), no. 4, pg. 527–546.
- [5] David H. Bailey and Michal Misiurewicz, “A strong hot spot theorem,” *Proceedings of the American Mathematical Society*, vol. 134 (2006), no. 9, pg. 2495–2501.
- [6] Veronica Becher and Santiago Figueira, “An example of a computable absolutely normal number,” *Theoretical Computer Science*, vol. 270 (2002), pg. 947–958.
- [7] Veronica Becher and Santiago Figueira, “Turing’s unpublished algorithm for normal numbers,” *Theoretical Computer Science*, vol. 377 (2007), pg. 126–138.
- [8] L. Berggren, J. M. Borwein and P. B. Borwein, *Pi: a Source Book*, Springer-Verlag, Third Edition, 2004.
- [9] Jonathan Borwein and David H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, AK Peters, Natick, MA, 2008.
- [10] Jonathan M. Borwein and Peter B. Borwein, “On the generating function of the integer part: $[n\alpha + \gamma]$,” *Journal of Number Theory*, vol. 43 (1993), pg. 293–318.
- [11] Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, Wiley-Interscience, New York, 1987.

- [12] Cristian S. Calude, “Borel normality and algorithmic randomness,” in G. Rozenberg, A. Salomaa, ed., *Developments in Language Theory*, World Scientific, Singapore, 1994, pg. 113–119.
- [13] Cristian S. Calude and Michael J. Dinneen, “Computing a glimpse of randomness,” *Experimental Mathematics*, vol. 11 (2002), pg. 361–370.
- [14] A. H. Copeland and P. Erdős, “Note on normal numbers,” *Bulletin of the American Mathematical Society*, vol. 52 (1946), pg. 857–860.
- [15] Peter Hertling, “Simply normal numbers to different bases,” *Journal of Universal Computer Science*, vol. 8 (2002), pg. 235–242.
- [16] Hajime Kaneko, “On normal numbers and powers of algebraic numbers,” *Integers*, vol. 10 (2010), pg. 31–64.
- [17] Davar Khoshnevisan, “Normal numbers are normal,” Clay Mathematics Institute, 2006, available at http://www.claymath.org/library/annual_report/ar2006/06report_normalnumbers.pdf.
- [18] Greg Martin, “Absolutely abnormal numbers,” *American Mathematical Monthly*, vol. 108 (October 2001), pg. 746–754.
- [19] R. Stoneham, “On absolute (j, ε) -normality in the rational fractions with applications to normal numbers,” *Acta Arithmetica*, vol. 22 (1973), pg. 277–286.

6 Appendix

Proof of Theorem 2 (Given coprime $b \geq 2$ and $c \geq 2$, $B = b^p c^q r$ for integers $p, q, r \geq 1$, neither b nor c divide r , and $D = c^{q/p} r^{1/p} / b^{c-1} < 1$, the Stoneham constant $\alpha_{b,c}$ is B -nonnormal):

Let $n = \lfloor c^m / p \rfloor$, and let $w = np / c^m$, so that $n = wc^m / p$. Note that for even moderately large m , relative to p , the fraction w is very close to one. Let Q_m be the shifted fraction of $\alpha_{b,c}$ immediately following position n in its base- B expansion. One can write

$$\begin{aligned} Q_m &= B^n \alpha_{b,c} \bmod 1 \\ &= \left(\sum_{k=0}^m b^{pn-c^k} c^{qn-k} r^n \right) \bmod 1 + \sum_{k=m+1}^{\infty} b^{pn-c^k} c^{qn-k} r^n \end{aligned} \quad (14)$$

$$= \sum_{k=m+1}^{\infty} b^{pn-c^k} c^{qn-k} r^n = \sum_{k=m+1}^{\infty} \frac{c^{qwc^m/p-k} r^{wc^m/p}}{b^{c^k-wc^m}}. \quad (15)$$

(The first summation in (14) vanishes because all summands are integers.) Thus Q_m is accurately approximated (in ratio) by the first term of the series (15), namely

$$S_1 = \frac{1}{c^{m+1}} \left(\frac{c^{qw/p} r^{w/p}}{b^{c-w}} \right)^{c^m}, \quad (16)$$

and this in turn is very accurately approximated (in ratio) by

$$S'_1 = \frac{D^{c^m}}{c^{m+1}}, \quad (17)$$

where $D = c^{q/p} r^{1/p} / b^{c-1}$ as defined in the hypothesis. So for all sufficiently large integers m ,

$$S'_1(1 - 1/10) < Q_m < S'_1(1 + 1/10). \quad (18)$$

Given that $D < 1$, as assumed in the hypothesis, it is clear from (17) that Q_m will be very small for even moderate-sized m , and thus the base- B expansion of $\alpha_{b,c}$ will feature long stretches of zeroes beginning immediately after position n , where $n = \lfloor c^m / p \rfloor$. In particular, given $m \geq 1$, let $Z_m = \lfloor \log_B 1/Q_m \rfloor$ be the number of zeroes that immediately follow position $\lfloor c^m / p \rfloor$. Then after noting that $B \geq 6$ (implied by the definition of b, c, p, q, r above), we can rewrite (18) as

$$c^m \log_B(1/D) + (m+1) \log_B c - 2 < Z_m < c^m \log_B(1/D) + (m+1) \log_B c + 2. \quad (19)$$

Now let F_m be the fraction of zeroes up to position $c^m + Z_m$. Clearly

$$F_m > \frac{\sum_{k=1}^m Z_k}{c^m + Z_m}, \quad (20)$$

since the numerator only counts zeroes in the long stretches, ignoring many others in the “random” stretches. The summation in the numerator satisfies

$$\begin{aligned} \sum_{k=1}^m Z_k &> \frac{c}{c-1} \left(c^m - \frac{1}{c} \right) \log_B(1/D) + \frac{m(m+3)}{2} \log_B c - 2m \\ &> \frac{c^{m+1}}{c-1} \left(1 - \frac{1}{c^{m+1}} \right) \log_B(1/D) - 2m. \end{aligned} \quad (21)$$

Thus given any $\varepsilon > 0$, we can write, for all sufficiently large m ,

$$\begin{aligned} F_m &> \frac{\frac{c^{m+1}}{c-1} \left(1 - \frac{1}{c^{m+1}} \right) \log_B(1/D) - 2m}{c^m + c^m \log_B(1/D) + (m+1) \log_B c} \\ &= \frac{\frac{c}{c-1} \left(1 - \frac{1}{c^{m+1}} \right) \log_B(1/D) - \frac{2m}{c^m}}{1 + \log_B(1/D) + \frac{m+1}{c^m} \log_B c} \\ &\geq \frac{\frac{c}{c-1} \log_B(1/D)}{1 + \log_B(1/D)} - \varepsilon = T - \varepsilon, \end{aligned} \quad (22)$$

where

$$T = \frac{c}{c-1} \cdot \frac{\log_B(1/D)}{1 + \log_B(1/D)}. \quad (23)$$

To prove our desired result, it suffices to establish that $F_m > T > 1/B$, which means that infinitely often (namely on segments up to position $c^m + Z_m$ for positive integers m) the fraction of zeroes exceeds the “normal” frequency of a zero, namely $1/B$, by the nonzero amount $T - 1/B$. But depending on the particular values of b, c, p, q and r , the condition $T > 1/B$ might not hold. Recall that the calculation above ignores the many zeroes in the “random” portions of the expansion, and thus the estimate T might not be sufficiently accurate to establish nonnormality, at least not in the single-digit frequency sense.

However, a simple modification of the above argument can establish nonnormality in the multi-digit frequency sense. Note that given any integer $M > 1$, then for all m with $Z_m > M$, we will see an M -long string of zeroes beginning immediately after position n , where $n = \lfloor c^m/p \rfloor$ as above. Indeed, the condition that an M -long string

of zeroes begins at position t will be fulfilled for $\bar{Z}_m = Z_m - M + 1$ consecutive positions beginning with $t = n + 1 = \lfloor c^m/p \rfloor + 1$. Note that for sufficiently large m , the modified count \bar{Z}_m is nearly as large as Z_m . What's more, when we sum \bar{Z}_k for $k = 1$ to m , we obtain, as in (21) above,

$$\begin{aligned} \sum_{k=1}^m \bar{Z}_k &> \frac{c}{c-1} \left(c^m - \frac{1}{c} \right) \log_B(1/D) + \frac{m(m+3)}{2} \log_B c - (M+1)m \\ &> \frac{c^{m+1}}{c-1} \left(1 - \frac{1}{c^{m+1}} \right) \log_B(1/D) - (M+1)m. \end{aligned} \quad (24)$$

But the small term $(M+1)m$ in this expression disappears when we divide by c^m and take the limit as in (22) above. Thus we obtain exactly the same limiting bound T as we calculated above in (23) for individual zeroes. Note that the natural frequency for an M -long string of zeroes is $1/B^M$. Since $T > 1/B^M$ for all sufficiently large M , we conclude that $\alpha_{b,c}$ is B -nonnormal. **QED**

Proof of Theorem 3 (If α_{b_1,c_1} and α_{b_2,c_2} are each B -nonnormal according to Theorem 2, then $\alpha_{b_1,c_1} + \alpha_{b_1,c_1}$ is B -nonnormal):

Given the hypothesized conditions, the proof of Theorem 2 established that the base- B expansion of α_{b_1,c_1} has long stretches of zeroes beginning at positions $P_{1,m} = \lfloor c_1^m/p_1 \rfloor + 1$ (for positive integers m), extending for length $Z_{1,m} \approx c_1^m \log_B(1/D_1) \approx P_{1,m} p_1 \log_B(1/D_1)$, where $D_1 = c_1^{q_1/p_1} r_1^{1/p_1} / b_1^{c_1-1}$. Similarly, the base- B expansion of α_{b_2,c_2} has long stretches of zeroes beginning at positions $P_{2,n} = \lfloor c_2^n/p_2 \rfloor + 1$ (for positive integers n), extending for length $Z_{2,n} \approx c_2^n \log_B(1/D_2) \approx P_{2,n} p_2 \log_B(1/D_2)$, where $D_2 = c_2^{q_2/p_2} r_2^{1/p_2} / b_2^{c_2-1}$. In each case, the approximation indicated is as accurate in ratio as desired, for all sufficiently large m or n , respectively.

Note that the base- B expansions of the two constants will share a long stretch of zeroes, provided there exists some pair of integers (m, n) such that the corresponding starting points $P_{1,m}$ and $P_{2,n}$ are very close in ratio. In that case, the corresponding strings of zeroes will overlap for a length L that is close in ratio to the shorter of the two lengths. In other words,

$$\begin{aligned} L &\approx \min(Z_{1,m}, Z_{2,n}) \approx \min(P_{1,m} p_1 \log_B(1/D_1), P_{2,n} p_2 \log_B(1/D_2)) \\ &\approx P_{1,m} \min(p_1 \log_B(1/D_1), p_2 \log_B(1/D_2)) = P_{1,m} E, \end{aligned} \quad (25)$$

where $E = \min(p_1 \log_B(1/D_1), p_2 \log_B(1/D_2))$, and where the approximations shown are as close in ratio as desired for all sufficiently large m and n .

What's more, since the base- B expansions of α_{b_1, c_1} and α_{b_2, c_2} share this section of zeroes, beginning at position $P_{1,m} \approx P_{2,n}$ and continuing for length L , so will the base- B expansion of $\alpha_{b_1, c_1} + \alpha_{b_2, c_2}$.

Now suppose that we can construct a sequence of pairs of integers (m_k, n_k) , where the above condition, namely $P_{1,m_k} \approx P_{2,n_k}$ culminating with $L_k \approx P_{1,m_k} E$, is met for each k . At each k , even if we count only the zeroes in the common stretch L_k (ignoring all zeroes in all stretches and all “random” segments that precede it), we obtain, as an estimate of the fraction F_k of zeroes up to position $P_{1,m_k} + L_k$,

$$F_k \geq \frac{L_k}{P_{1,m_k} + L_k} \approx \frac{P_{1,m_k} E}{P_{1,m_k} + P_{1,m_k} E} = \frac{E}{1 + E}, \quad (26)$$

where the approximation is as accurate as desired (in absolute terms, not just in ratio) for all sufficiently large k . Recall that $E = \min(p_1 \log_B(1/D_1), p_2 \log_B(1/D_2)) > 0$ by hypothesis, so that that the expression $E/(1 + E)$ is independent of k and strictly greater than zero.

Such a sequence of integer pairs (m_k, n_k) can be constructed as follows: First consider the simpler special case where $p_1 = p_2$. Given $\epsilon > 0$, we require that for all sufficiently large pairs (m_k, n_k) ,

$$1 - \epsilon < \frac{P_{1,m_k}}{P_{2,n_k}} < 1 + \epsilon. \quad (27)$$

But this can equivalently be rewritten in any of the forms

$$\begin{aligned} 1 - \epsilon &< \frac{c_1^{m_k}}{c_2^{n_k}} < 1 + \epsilon \\ -\epsilon &< m_k \log c_1 - n_k \log c_2 < \epsilon \\ \left| \frac{m_k}{n_k} - \frac{\log c_2}{\log c_1} \right| &< \frac{\epsilon}{n_k \log c_1}. \end{aligned} \quad (28)$$

This last condition is fulfilled if we specify, for the sequence of pairs (m_k, n_k) , the sequence of fractions produced by the continued fraction approximation for $\log c_2 / \log c_1$. Recall that the error in the continued fraction approximation at each step is less than the square of the reciprocal of the current denominator [11, pg. 373]. Thus we can write,

$$\left| \frac{m_k}{n_k} - \frac{\log c_2}{\log c_1} \right| < \frac{1}{n_k^2} < \frac{\epsilon}{n_k \log c_1}, \quad (29)$$

for all sufficiently large k , which satisfies the condition in (28) (and thus in 27 also).

Now consider the more general case where p_1 is not necessarily the same as p_2 . Given $\epsilon > 0$, we require that for all sufficiently large pairs (m_k, n_k) ,

$$\begin{aligned}
1 - \epsilon &< \frac{P_{1,m_k}}{P_{2,n_k}} < 1 + \epsilon \\
1 - \epsilon &< \frac{c_1^{m_k}/p_1}{c_2^{n_k}/p_2} < 1 + \epsilon \\
-\epsilon &< m_k \log c_1 - n_k \log c_2 + (\log p_2 - \log p_1) < \epsilon.
\end{aligned} \tag{30}$$

In this case, we can apply a generalization of the continued fraction algorithm presented in [10] (Algorithm 0.3) to construct the requisite sequence of integer pairs (m_k, n_k) . A simple normalization of (30) reduces it to the form required in [10].

In short, for any choice of coprime pairs of integers (b_1, c_1) and (b_2, c_2) satisfying the hypothesis, we can construct an infinite sequence of positions $(P_{1,m_k} + L_k)$ in the base- B expansion of $\alpha_{b_1,c_1} + \alpha_{b_2,c_2}$ such that the fraction F_k of zeroes up to position $P_{1,m_k} + L_k$ exceeds the fixed bound $E/(1 + E)$. If this bound satisfies $E/(1 + E) > 1/B$, we are done. If not, a simple extension of the preceding argument to count the number of indices where an M -long strings of zeroes begins, as was done near the end of the proof of Theorem 2, shows that the asymptotic bound $E/(1 + E)$ also applies to the frequency of M -long strings of zeroes. Since for all sufficiently large M , the condition $E/(1 + E) > 1/B^M$ is satisfied, this concludes the proof.

QED